

(TM)

Distribution rights by Intellegenetics, Inc.

- n.a. database search, using Smith-Waterman algorithm

ot generated

GGGGTCGGCGCGGAGGT.....AAGTGCCTTTT

TABLE default

Database 0; Que

362067 seqs, 549138275 bases x 2

Listing first 45 summaries

emb1-new3

genbank99

99:VRL4 :

122:part1

Mean 11.533; Variance 4.446; scale 2.594

ved by analysis of the total score distribution.

9

1111

REFERENCE

AUTHORS

5711

ПОПРАВКА:

ΠΥΛΩΝ

FEATURES

5017

C

BASE COUNT

ORIGIN

[illegible]

ORGANISM	REFERENCE	TITLE	JOURNAL	FEATURES	CDS	BASE COUNT	ORIGIN
Parasitus auratus	1 (bases 1 to 1933)	Glasgow, E. and Schechter, N.	Nucleotide sequence of a GFAP - like intermediate filament cDNA from goldfish retina	Unpublished (1993)			
	1. 1933		Location/Qualifiers				
	/organism="Parasitus auratus"						
	/dev_stage="adult"						
	/tissue_type="retina"						
	20..1099						
	/gene="GFAP-1"						
	/note="putative"						
	/codon_start=1						
	/db_xref="PID:g388623"						
	/translation="MGINDRFASYLEKRELEQNKMLVAELNLRGKERSRLDIOV EELERLROVDGLNAGKARLEIEDNINSLATLTKOJENALROAEKNILNTPRODD VDEALNRVOLDERKIDALOEISPLRVEHEEMKLOEOLIAQOVDIVSKPLDTT AKKIRAPFAMATSNNOETFEWRSFALITDAAGNARLRANKOAEADYRPOLOG LTCDESLSRGSNESLEKQLEMERERIRIETAGYDVIARLEDELOMKREMARLOET ODLNLVRLADLEIATYRKLEGESESITVQNFNTLQFRDTSLDTRKPEAVKRS IVATVETRGEIETKESITERKDLF"						
	561 a 395 c 507 g 470 t						
	Query Match 1.6%; Score 25; DB 49; Length 1933;						
	Best Local Similarity 77.8%; Pred. No. 3.26e-01;						
	Matches 35; Conservative 0; Mismatches 10; Indels 0; Gaps 0;						
Db 901	gaagctgctcgaagagagaagacagatcttcctgcgtgca 945						
Cp 931	gaagctgctcgaagagagtgaggttgacatgccatgctgca 887						
RESULT 9	RRMAP1B5 7095 bp RNA ROD 21-OCT-1992						
LOCUS	R.norvegicus mRNA for microtubule associated protein IB.						
DEFINITION	X60370 X60371 X60550						
ACCESSION	957618						
NID	MAP1B gene; microtubule-associated protein.						
KEYWORDS	Norway rat.						
SOURCE	Rattus norvegicus						
ORGANISM	Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Euteiria; Rodentia; Sciurognathi; Myomorphae; Muridae; Murinae; Rattus.						
REFERENCE	1 (bases 1 to 7095)						
AUTHORS	zauner, W., Kratz, J., Staunton, J., Feick, P. and Wiche, G.						
TITLE	Identification of two distinct microtubule binding domains on recombinant rat MAP 1B						
JOURNAL	Eur. J. Cell Biol. 57 (1), 66-74 (1992)						
MEDLINE	92347374						
REFERENCE	2 (bases 1 to 7095)						
AUTHORS	Wiche, G.						
TITLE	Direct Submission						
JOURNAL	Submitted (07-NOV-1991) G. Wiche, Inst of Biochemistry, University of Vienna, Waehringerstrasse 17, 1090 Vienna, AUSTRIA						
REMARK	revised by (3)						
REFERENCE	3 (bases 1 to 7095)						
AUTHORS	Wiche, G.						
TITLE	Direct Submission						
JOURNAL	Submitted (07-NOV-1992) G. Wiche, Institute of Biochemistry and Molecular Biology, University of Vienna, Dr. Bohrergasse 9, 1030 Vienna, AUSTRIA						
FEATURES	Location/Qualifiers						
SOURCE	1..7095						
	/organism="Rattus norvegicus"						
	/strain="Sprague-Dawley"						
	/dev_stage="adult"						

```

/tissue_type="brain"
/cell_type="C6 glioma"
BASE COUNT      2124 a 1856 c 1799 g 1316 t
ORIGIN
Query Match      1.6%; Score 24; DB 92; Length 7095;
Best Local Similarity 76.1%; Pred. No. 1.38e+00;
Matches 35; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Db 1004 gcaaggaatcagctatttcacagcagctgagctggaaccacaa 1049
||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 305 GCAAGGAGCTGAGTACGTCAGACGAGAGTGCATGCAACCCACAA 350

RESULT 10
LOCUS      PCHSAB      7322 bp      DNA      PLN      06-JUL-1994
DEFINITION P.sativum gene for chalcone synthase.
ACCESSION  X80007
            G510542
ORIGIN
ORGANISM   Pisum sativum
            pea.
            Eukaryote; mitochondrial eukaryotes; Viridiplantae;
            Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;
            Magnoliopsida; Rutanae; Sapindales; Fabaceae; Papilionoideae;
            Pisum.
            1 (bases 1 to 7322)
REFERENCE   1
AUTHORS    Hellens,R.P.
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 7322)
AUTHORS    Hellens,R.P.
TITLE      Direct Submission
JOURNAL     Submitted (01-JUL-1994) R.P. Hellens, John Innes Institute, Colney
            Lane, Norwich NR4 7DH, UK
FEATURES
source      1..7322
            /organism="Pisum sativum"
            /strain="J1 813"
            /clone="1b="1abda GEM-11"
            /clone="B"
            /map="linkage group 3"
            <1040..1217
            /gene="CHS-1A"
            /number=1
            join(<1040..1217,1328..>2319)
            /gene="CHS-1A"
            join(1040..1217,1328..2319)
            /gene="CHS-1A"
            /EC_number="2.3.1.74"
            /codon_start=1
            /product="naringenin-chalcone synthase"
            /db_xref="PID:G510543"
            /translation="MTVNEIRAOAKBEGPATVFAIGTTPONCYEOSTPYDFPRIT
            NSOHTLEKREKFORCDSKMTKRMMHLEELIKENPSCIEFMAESIDAROMYVEV
            PKLGEATKAIKENGQPSKTIHLICTTSVDPMGADYQTLKLGRLPYKRMV
            OOGCPAGGDPLELADLAENNGANLVVCSITVTEFGPDDTHLDSLVGALFGDG
            AAAYVGSDDLPELVEKPLFELVTAQTIIVDESGAIDGRLRAGLFLHLKRPGLIS
            KNIERALVEAFQPLGIDYNSLFWIAHPGPAIILDOVEAKLGIKORKMQRHVSSEY
            GMSACVLEFIIDEMRRSKEDGLATGTGGLMGVLFGRGPDLTETVYVLSHMAI"
            1218..1327
            /gene="CHS-1A"
            /number=1
            join(1328..>2319)
            /gene="CHS-1A"
            /number=2
            join(<4336..4513,4642..>5633)
            /gene="CHS-1B"
            join(4336..4513,4642..5633)
            /gene="CHS-1B"
            /EC_number="2.3.1.74"
            /codon_start=1
            /product="naringenin-chalcone synthase"
            /db_xref="PID:G510544"

```

```

/translation="MTVNEIRAOAKBEGPATVFAIGTTPONCYEOSTPYDFPRIT
NSEHTLEKREKFORCDSKMTKRMMHLEELIKENPSCIEFMAESIDAROMYVEV
PKLGEATKAIKENGQPSKTIHLICTTSVDPMGADYQTLKLGRLPYKRMV
OOGCPAGGDPLELADLAENNGANLVVCSITVTEFGPDDTHLDSLVGALFGDG
AAAYVGSDDLPELVEKPLFELVTAQTIIVDESGAIDGRLRAGLFLHLKRPGLIS
KNIERALVEAFQPLGIDYNSLFWIAHPGPAIILDOVEAKLGIKORKMQRHVSSEY
GMSACVLEFIIDEMRRSKEDGLATGTGGLMGVLFGRGPDLTETVYVLSHMAI"
<4336..4513
/gene="CHS-1B"
/number=1
4514..4641
/gene="CHS-1B"
/number=1
4642..>5633
/gene="CHS-1B"
/number=2
BASE COUNT      2350 a 1307 c 1264 g 2401 t
ORIGIN
Query Match      1.6%; Score 24; DB 64; Length 7322;
Best Local Similarity 81.6%; Pred. No. 1.38e+00;
Matches 31; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Db 654 aaagcgtaccctctatggaataaccacaaatgatt 691
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1404 AAAGACTAACATCTCTCCATTAACCCCAATAGGTT 1441

RESULT 11
LOCUS      CEC29F3      18632 bp      DNA      INV      31-JAN-1997
DEFINITION Caenorhabditis elegans cosmid C29F3.
ACCESSION  Z81043
            G1627634
KEYWORDS   Caenorhabditis elegans.
SOURCE     Eukaryote; mitochondrial eukaryotes; Metazoa; Nematoda;
            Secernentea; Rhabdita; Rhabditida; Rhabditina; Rhabditodea;
            Rhabditidae; Peloderinae; Caenorhabditis.
            1 (bases 1 to 18632)
REFERENCE   1
AUTHORS    Matthews,L.
JOURNAL     Direct Submission
            Submitted (21-OCT-1996) Nematode Sequencing Project, Sanger Centre,
            Hinxton, Cambridge CB10 1RQ, England and Department of Genetics,
            Washington University, St. Louis, MO 63110, USA. E-mail:
            jesssanger.ac.uk or twenematode.wustl.edu
            2 (bases 1 to 18632)
REFERENCE   2
AUTHORS    Wilson,R., Almscough,R., Anderson,K., Baynes,C., Berks,M.,
            Bonfield,J., Burton,J., Connell,M., Copsey,T., Cooper,J.,
            Coulson,A., Craxton,M., Dear,S., Du,Z., Durbin,R., Favello,A.,
            Fulston,L., Gardner,A., Green,P., Hawkins,T., Hillier,L., Jier,M.,
            Johnston,L., Jones,M., Kershaw,J., Kirsten,J., Lalster,N.,
            Latelle,P., Lightning,J., Lloyd,C., McMurtry,A., Mortimore,B.,
            O'Callaghan,M., Parsons,J., Percy,C., Ritken,L., Roopra,A.,
            Saunders,D., Showkeen,R., Smaildon,C., Smith,A., Sonhammer,E.,
            Staden,R., Sulston,J., Thierry-Mieg,J., Thomas,K., Vaadin,M.,
            Vaughan,K., Waterston,R., Watson,A., Weinstock,L.,
            Wilkison-Spratt,J. and Woldman,P.
            2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
            elegans
            Nature 368 (6466), 32-38 (1994)
            94150718
COMMENT
JOURNAL     MEDLINE
            94150718

```

Current sequence finishing criteria for the *C. elegans* genome sequencing consortium are that all bases are either sequenced unambiguously on both strands, or on a single strand with both a dye primer and dye terminator reaction, from distinct subclones. Exceptions are indicated by an explicit note. IMPORTANT: This sequence is NOT necessarily the entire insert of clone C29F3. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions. Coding sequences below are predicted from computer analysis, using the program Genefinder (P. Green, ms in preparation), and other

available information.

FEATURES
SOURCELocation/Qualifiers
1.18632

/organism="Caenorhabditis elegans"

/chromosome="V"

complement(join(158..344,421..598,647..698,743..796,
845..1128,1176..1414,1456..1544,1591..1788,2036..2201,
2721..2787))

/note="protein predicted using GeneFinder"

/codon_start=1

/db_xref="PID:e299269"

/product="C29F3.d"

/db_xref="PID:g1813905"

/translation="MTPADNRYVSNALAEFVGEHSRRRRRMTSLQYGLARPD
KTIISFRHCHRVETGCGRIANRFLPTQLTSSMTLLVSGLSLWQDVESRCG
ESTIPFLEILSGHVGARPTCFGHPGQYLPPTAKTSILNRLLDGLRDSLE
TYPEPDSKIKYQNSCEPQSCDSKDWVGGEPTDAFOVAYOCCTYARL
RESIDRIATVAGEIYIGEVYQNSQDSYDINTEKSDENGVEYEVNIRRA
LDPNARRIDEVSESENTIRKVGOKPIAQANVAVNAPTEAGTEDGVEDQTVY
IEETIAOGGFEVNETTAPRACGPFOAGPERRGAGQGPRAQGPPOQFPQNC
PVYQDPVQFPAPVGIAPAPAGMQLHCFPADLVNVEYERKDELEGVDMQAL
HGKETTSPVAKYWLHDPEDQAEFEVLENGESFTLEKLVFATDQOQVKKLDDL
NPTSTGKINGCEFFMAQPEVRSILCY"complement(join(5522..5692,5740..5834,5888..6062,
6642..6740))

/note="protein predicted using GeneFinder"

/codon_start=1

/product="C29F3.c"

/db_xref="PID:e299497"

/db_xref="PID:g1813904"

/translation="MTIIILFAFFCMSEVSGCIPMTPEEPVYVYVPCPAGWESFO
RANGIMCYIAKPPACWTSPOQYCONQMSPTFNGESAERQFQIDALAAHLAS
QFLHGIIRSCAAPCATPDPVWQNGSNDNNPANDYIQYDSSGGLSMDLAKGE
INDITCALIRVSCGYAA"complement(join(7079..7249,7299..7387,7433..7601,
8197..8292))

/note="protein predicted using GeneFinder"

/codon_start=1

/product="C29F3.b"

/db_xref="PID:e299268"

/db_xref="PID:g1813903"

/translation="MMWRLTYLFFASAOISGCLPMYPPPEPVYVPCPAGWESFO
ATGIMCYIPATPGAGMTPOAACQANGANGESAERQFQIDALAAHLAS
HIGMRCARCTVNDPFWMLNGVNDNTFANDYISLDLGDCLSMGLGNGQYNDIT
CDATFAYSCGPPAA"complement(join(10724..10982,11039..11394,12940..13052,13183..13295,
13344..13893,13948..14262,14952..15141))

/note="protein predicted using GeneFinder"

/codon_start=1

/product="C29F3.g"

/db_xref="PID:e299270"

/db_xref="PID:g1813906"

/translation="MYPEPSTFFEDHSRDVDFEQENVEIFGLVAVFHDVSLFFIIS
LNRIISVCMPLFYTKFMNLKYTKLFIFAVWLISLIGLSLIVFCIRIPNADLIFPI
ILASVCEIEGVDLFRNSCIYVIMMLIVETLIXKILSNTEKKSQAOVQFS
RDRFLQATYQGSILFLATVFFPTFSVNSWTFIIGNSCGRWIVPQYFATIL
TNITSSNNSAPVQAVQWQWQVQVSTSESYSESYNGENIKLSGNRNVQGFSL
FMDYQFIOIKYLFNSATLPLYSNMPDQVITADTRISATIVPPGVDSQYSLR
GVLEFDGTNRSTYGLTGLTGSKQSYSTISYSDSNVANSNPVADYEN
TKSICSEMIICPDQDCGIVSLDGTGSAALQYLNKPEFSLVANKTKGFGLEAF
DGVYTKNLITTYVNRSENLPOEFGRLLTVLFETISALITTRDDADAFDETTS
RKGITSDSTGNSQEAFLINAPVKNEDYIFKVAIADLGSGLTLVLCRYTELAR
ELSEFNSNPLPLNTTFHGNFLHIEYIPNTDSKLFLEFDEIKSVTSRSEFFACV
IVFLFLF"complement(join(15422..15713,16068..16591,16640..17140,
17376..18179))

/note="protein predicted using GeneFinder"

/codon_start=1

/product="C29F3.a"

/db_xref="PID:e299496"

/db_xref="PID:g1813902"

/translation="MNETLDVRVRLKRLFLFSYKQFQSDQSVKAIIVMSGKPNSEVAG
IVFLFLF"complement(join(15422..15713,16068..16591,16640..17140,
17376..18179))

/note="protein predicted using GeneFinder"

/codon_start=1

/product="C29F3.a"

/db_xref="PID:e299496"

/db_xref="PID:g1813902"

/translation="MNETLDVRVRLKRLFLFSYKQFQSDQSVKAIIVMSGKPNSEVAG
IVFLFLF"

ADIDMEKAEKTAAGVSNLREBQOKOLITELISOKPIYAAIMSGMGGLEFIALACHR
IAYDKRTILGEPVTLGIMPGDQGRTRKRETFYOVVLDLTLGRIKANKAKTIT
VDKRTIPLGICITSTETTKYIEELIYQASARELANAKLITNRKGVHNQAVMTS
KFLVDNVLIMARKNLILITNGYPAKLLDVVRRYADPKNGVEAEAFGLSOT
FOSKALGLEGGVSTANAKKRYGKGLPVNEIAYVAGMGAGIANVITNKIRVILDA
NOAGVERGONHVAATHLNROKROKISLESEKIYNHLVPTIDYSAMNADVLEAVE
DLPIKHYVIOIENVGPNTLIASNTSALIKIDIAASSSDVIGWYHSPVEKKML
LEITTHGCTKEITLATAOIGIKOGKLYVVKCCPGFFVRCISPMKSEIVRLQEGV
EPSELDTLTRKFGPPVGAATLADENAGIDVAEHARYLGLALGPRHGSVDLSEIVR
AGHGRKTSIGITVYGDGASKVKNQENAKLEKYLPTIKSVSPEDRQLRVISF
VNEALICLEGVSSPSGDIVASVTFGFPFPGPRFVDLYGAGKLVAAIMRFAVY
YESVQFEPCOLLNRHAKSGGQFS"

BASE COUNT 5872 a 3349 c 3316 g 6095 t
ORIGIN
Query Match 1.6%; Score 24; DB 35; Length 18632;
Best Local Similarity 90.0%; Pred. No. 1.38e+00;
Matches 27; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 13078 ggaattaaagaactatgatgtggaataa 13107
Qy 1377 ggtactaaagaactatgatgtggaataa 1406

RESULT 12
LOCUS CER09D4 47745 bp DNA INV 06-NOV-1996
DEFINITION Caenorhabditis elegans cosmid T09D4.
ACCESSION Z81590
NID g1666006
KEYWORDS HTGS; HTGS_PHASE1.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans.
Eukaryote; mitochondrial eukaryotes; Metazoa; Nematoda;
Secernentea; Rhabdita; Rhabditida; Rhabditiina; Rhabditoidea;
Rhabditidae; Caenorhabditis.
1 (bases 1 to 47745)
AUTHORS Unknown.
TITLE Direct Submission
JOURNAL Submitted (04-NOV-1996) Nematode Sequencing Project, Sanger Centre,
Hinxton, Cambridge CB10 1RQ, England and Department of Genetics,
Washington University, St. Louis, MO 63110, USA. E-mail:
joesanger.ac.uk or rwenematode.wustl.edu
COMMENT of segments is not known; 800 nt's separate segments.
Cosmid-T09D4; Contig ID=00659; Length=42167; Status=Unfinished
Cosmid-T09D4; Contig ID=01673; Length=41019; Status=Unfinished
Cosmid-T09D4; Contig ID=01596; Length=1665; Status=Unfinished
Cosmid-T09D4; Contig ID=01600; Length=1271; Status=Unfinished.

FEATURES
source location/Qualifiers
1.47745
/organism="Caenorhabditis elegans"
/clone="T09D4"
/chromosome="V"

BASE COUNT 14131 a 8322 c 8318 g 14134 t 2840 others
ORIGIN
Query Match 1.6%; Score 24; DB 38; Length 47745;
Best Local Similarity 90.0%; Pred. No. 1.38e+00;
Matches 27; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 32078 ttcttcacatcaatgtcttcaattcc 32107
Cp 1406 ttcttcacatcaatgtcttcaattcc 1377

RESULT 13
LOCUS HS799F10 115419 bp DNA HTG 13-JAN-1997
DEFINITION Human DNA sequence *** SEQUENCING IN PROGRESS *** from clone
799F10; HTGS phase 1.
ACCESSION Z82245
NID g1780984
KEYWORDS HTG; HTGS_PHASE1.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL

Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Euthera; Primates; Catarrhini; Homidae; Homo. 1 (bases 1 to 115419)

COMMENT

Direct Submission
Submitted (06-NOV-1996) Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humbry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk IMPORTANT: This sequence is unfinished and does not necessarily represent the correct sequence. Work on the sequence is in progress and the release of this data is based on the understanding that the sequence may change as work continues. The sequence may be contaminated with foreign sequence from E.coli, yeast, vector, phage etc. Order of segments is not known; 800 n's separate segments. Unfinished sequence: BK799910 Contig_ID: 00052 Length: 992 bp Unfinished sequence: BK799910 Contig_ID: 00061 Length: 885 bp Unfinished sequence: BK799910 Contig_ID: 00185 Length: 12614 bp Unfinished sequence: BK799910 Contig_ID: 00949 Length: 830 bp Unfinished sequence: BK799910 Contig_ID: 00953 Length: 785 bp Unfinished sequence: BK799910 Contig_ID: 00960 Length: 958 bp Unfinished sequence: BK799910 Contig_ID: 01506 Length: 6932 bp Unfinished sequence: BK799910 Contig_ID: 01718 Length: 38510 bp Unfinished sequence: BK799910 Contig_ID: 02165 Length: 13742 bp Unfinished sequence: BK799910 Contig_ID: 02195 Length: 2234 bp Unfinished sequence: BK799910 Contig_ID: 02223 Length: 1000 bp Unfinished sequence: BK799910 Contig_ID: 02225 Length: 1341 bp Unfinished sequence: BK799910 Contig_ID: 02226 Length: 1016 bp Unfinished sequence: BK799910 Contig_ID: 02214 Length: 19570 bp.

FEATURES
SOURCE

*** WARNING: Phase 1 High Throughput Genome Sequence ***
* This sequence is unfinished. When sequencing is complete,
* the sequence data presented in this record will be replaced
* by a single finished sequence with the same accession number.
Location/Qualifiers
1..115419
/organism="Homo sapiens"
/clone="799910"
/chromosome="22"
24169 a 28425 c 26764 g 23990 t 12071 others

BASE COUNT 24169 a 28425 c 26764 g 23990 t 12071 others

ORIGIN

Query Match 1.6% Score 24; DB 34; Length 115419;
1st Local Similarity 80.0%; Pred. No. 1.38e+00;
Matches 32; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Db 109739 tcttctgttctgttcttccactctcgaagcgaag 109778
||||| ||||| ||||| ||||| ||||| |||||
Cp 833 TCTTGTGTTGATGTTTCATTAACCTTCAGACCTGGAAG 794

RESULT 14
LOCUS D90904 150894 bp DNA BCT 20-NOV-1996
DEFINITION Synechocystis sp. PCC6803 complete genome, 6/27, 630555-781448.
ACCESSION D90904
KEYWORDS g1652225
formaldehyde dehydrogenase; ABC transporter; DNA ligase; DNA polymerase III alpha subunit; Gumb protein; lysr transcriptional regulator; Mg-protoporhyrin IX; N-acetylornithine aminotransferase; NADH dehydrogenase subunit 4; NADH dehydrogenase subunit 5; acetate kinase; aspartacylase(ASP); cation-transporting ATPase; cell division protein FtsH; chemotaxis protein cheA; cytochrome oxidase d subunit I; cytochrome oxidase d subunit II; dieneolactone hydroxylase; dihydroflavonol 4-reductase; dihydropterocate pyrophosphorylase; esterase; ferrous iron transport protein B; fibrillin; formaldehyde dehydrogenase (glutathione); high light-inducible protein; leader peptidase I; lysostaphin; malonyl coenzyme A-acyl carrier protein transacylase; methyl-accepting chemotaxis protein (MCP); oligopeptide transport ATP-binding

SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL

protein OPP; oxygen independent coproporphyrinogen III oxidase; penicillin-binding protein 1b; phenoxymethylbenzate dihydroxylase; phycoerythrin b subunit; phycoerythrin associated linker protein; phycoerythrin b subunit; potassium channel; protein conferring resistance to acetazolamide, zam; regulatory component of sensory transduction system; regulatory components of sensory transduction system; ribonuclease II; sensory transduction histidine kinase; serine esterase; serine/threonine protein kinase; seryl-tRNA synthetase; sporulation protein SpoIID; sulfolipid biosynthesis protein SgdB; tRNA-Arg; tRNA-Leu; tyrosyl tRNA synthetase. Synechocystis sp. (strain:PCC6803) DNA.

REFERENCE
AUTHORS

1 (bases 1 to 150894)
Tabata, S.
Direct Submission
Submitted (28-JUN-1996) to the DDBJ/EMBL/GenBank databases. Satoshi Tabata, Kazusa DNA Research Institute, Laboratory of Gene Structure 2, 1573-3, Yanauchino, Kisarazu, Chiba 292, Japan
(E-mail: tabata@kazusa.or.jp, Tel: +81-438-52-3933, Fax: +81-438-52-3934)
2 (sites)
Kaneko, T., Sato, S., Kotani, H., Tanaka, A., Asamizu, E., Nakamura, Y., Miyajima, N., Hirosewa, M., Sugiyama, M., Sasamoto, S., Kimura, T., Hosouchi, T., Matsuno, A., Muraki, A., Nakazaki, N., Naito, K., Okumura, S., Shimpo, S., Takeuchi, C., Wada, T., Watanabe, A., Yamada, M., Yasuda, M. and Tabata, S.
Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions

JOURNAL
MEDLINE

97061201
Potential protein coding regions were assigned on the basis of similarity search of the ORFs and Genemark analysis.

FEATURES
SOURCE

Location/Qualifiers
1..150894
/organism="Synechocystis sp."
/strain="PCC6803"
complement(1..1038)
/note="ORF_ID:s110998"
/codon_start=1
/product="lysr transcriptional regulator"
/db_xref="PID:g1652226"
/transl_table=1
/translation="MSDIPFTLDLRILKAIASGSEFKRAADLYVSPAYSLQVONT
EKQSVPLFDGGRKAQLEAGHLLNNGEITILCOETCAIEDLOVLOGSTLVGA
SOTGTGTLPRMIGMFGROQPDVAVOVHSTRRTAVANGVDLAIIGGEVPELO
ETIVVLPRAEDBALILPVLHPLAKAETIQKEDYKLFISDSOSTIRKIDKYLQ
GEIDTRKLTIMELNSIETAKNAVOSGGAFAVSTATKELENNVLTADIKVEIR
RVLSVILPNPNRYRSKASNAFIREVLPOFSTHPDALDPERLPANPYSSNNRGDGGKD
GKSGSIEDSVY"
complement(1342..1539)
/note="ORF_ID:s11911"
/codon_start=1
/product="hypothetical protein"
/db_xref="PID:g1652227"
/transl_table=1
/translation="MSTQOQARALMMRHQFIKNRQSGMLSRAAAEIGVEAKDFWT
VQGRPOSSFFRTYDRSNASLS"
1659..2603
/note="ORF_ID:s11019"
/codon_start=1
/product="hypothetical protein"
/db_xref="PID:g1652228"
/transl_table=1
/translation="MNSPVAPMRYGFTLVVFTDQLGCGNPVAVFPAGEITDQGM
KIAEINSEVFLPVPTETGNGRRLRFTKRELDGHPITGTYLLGLQPSPL
VTTWQLEAPYAGLVPLVLYHKGQGVLTETLVAAQPKKDSAPSCEDALILGSDID
LQGEYEPQAYSCGLPFLPIPLINEALNLSFNFSVWNLACOMACVCYCLAAGDP
SLGSDNKLINGRPAPELGIAPEDATSGVAAGGYDDRLDPGSHHWQIEGKAL
GRPSLOITVVKDQGITAAVAVARSVYVSGFMNLGN"
2750..2830
/product="tRNA-Leu"

tRNA

CDS
 complement(3235..4326)
 /note="ORF_ID:s110997"
 /codon_start=1
 /product="hypothetical protein"
 /db_xref="PID:g1652232"
 /transl_table=11
 /translation="MAPKNSCHNGDMKQNSNLGEMIVAPYQSFHIGLLGLALASW
 /PACALIPSPPEPAEONQVNGDROEQVLAQNSDPTPTPEFDPSNDEKRLNGYKSG
 DVSASISYROEKHGDYESMEPTVTYKFKKIKYATLIGSSSPMIALLOTDDMDPNPY
 PAULATYATGACNCEYKFTSNONGSMYSRVDGFRNGGPHPADLNDMDNDPEVE
 VDNRLILFSTAGSHAPQALWALQGVYVDSFESFQFIRENNQSENKLPETIVA
 OSEKNGFLAAVANKALIGLEDMQTLKYYDRSDGLNCLLEYDQSNCLNEVK
 YDSPDALRAFLVEAGYIEAKFEAEENNTL"
 complement(4338..5693)
 /note="ORF_ID:s110996"
 /codon_start=1
 /product="hypothetical protein"
 /db_xref="PID:g1652230"
 /transl_table=11
 /translation="MNVSHRYHITIFGCGMKNKADSERMAGILENLTGMYTDDPNQAD
 LVMYTCSTRDAEOKVYSYLGROAKRKQREPELTVYAGCYAQOEGEQLLRRVELD
 LVMQPHANRLDQLLEQVWAGSQVYATESLHIMEDITKRRESIVSAMVNTITGCGER
 CICYVYVNRGVEQSRTPALIGEMEVLAQGFKEVTLTGQNDIAYGRDLPETPSGR
 BLHPTLDLYHYHDIIEGIDRLRPAISHRPYQERLQACQELPKYCEHFIPIFQSGDN
 DLKMKRGYTRKYLQILEKIRRYMPDAISADYVGFETEAQFENTLILIEVG
 FDLNTAAYSPPGVTPEAFVWQNLSEEVKQDRLNLNHLVSTQAMERSORYLGRVEV
 LVEGENLKSPPQVGMRTKGNRLTFQGEISELGTVPYKITEAPAFSLTGALSLVT
 A"
 5618..6969
 /gene="sqdb"
 /note="ORF_ID:s1r1020"
 /codon_start=1
 /product="sulfolipid biosynthesis protein sqdb"
 /db_xref="PID:g1652231"
 /transl_table=11
 /translation="MRALVIGGGGYCGNATALYLSNKGYEVLIDSLVRYWDAOLGA
 ETLPIAPIRORLDREVELTKIDLEIDINDYFELTNALROFOPDAVHGEORSA
 PPSMTDRBAVLTQANVNLNLALYALKEDPCHIKVIGMGEGYIPNDIEGYSI
 TIEHGRKDTLPYKPGSGFYHLISKVYSHNHFKKIKGLATLNLGIVGVTTEE
 TGMDEMLNRLDYDGVFGTALNRCFICQAIHPLVYVGGGTGLDLDRVRIETEL
 AIANPADRGQFERNQYTELFESVGDLAQWVQAGADLGKVIDHLENPRVLEBHYF
 NAVNTNLDLGLQHPFELSDLSLNLNFATKYKRDVQKHILPKVTWKG"
 7112..7402
 /note="ORF_ID:ssr1698"
 /codon_start=1
 /product="hypothetical protein"
 /db_xref="PID:g1652232"
 /transl_table=11
 /translation="MADPLTPAISDRICKHNNEDHASAIALYAYFGGQOTDVTNAQMO
 AIDPTGMLVYSESGSKTIRIEFEQPLKDSDAQVLIANAQKASVSKNS"
 complement(7439..8134)
 /note="ORF_ID:s110995"
 /codon_start=1
 /product="hypothetical protein"
 /db_xref="PID:g1652233"
 /transl_table=11
 /translation="MTVALDRRIYPPDSGQPMADNTEQFENVILKENLECFANHP
 DVFVGGDLMTVEBSGPEIRVADYVALGRKGRKSGROMQENNOAPQVVEFLSP
 GNTLKEMTKRLKFDHGVVEEYVDPDNLNETQVQRIGGETIIEENAHVWSPLLGI
 KFEELSAETLRVYVDPGRPELSTVALATAEQASCRANEAEQAREKIKRAELAEAND
 RLKALLAAGIDV"
 complement(8220..9029)
 /note="ORF_ID:s110994"
 /codon_start=1
 /product="hypothetical protein"
 /db_xref="PID:g1652234"
 /transl_table=11
 /translation="MSTNFPYPSIOHGYVOQKTVISRTNYIMLFHLANLTYWLLG
 LGIICLGMITISGDEDDLDVMSLEASSDVDTLDELVDGGMGEAEAPMLQY
 LSFELGVPLMILIGIDFSLMGVILNLTVAVTGIMPELNGAGIIFVLSLAI
 SLWGLASRLPAHLKFTESQSVSTERYIGCGIVTSKRLPYLANGTIGQAHVYDNAG
 NLITISVLPDPAVYIPHNQEIILIIDSPKGYGLAIKAKSDSDKWLKS"
 complement(8962..10059)

CDS
 /note="ORF_ID:s110993"
 /codon_start=1
 /product="potassium channel"
 /db_xref="PID:g1652235"
 /transl_table=11
 /translation="MGICSSQSEENLNLIDROGRRLROELMAGATTLAGLFVYVGTAMY
 RYEDMTWMDAFYMTTITLAVFGFETPLPSASLEFILLIMDLITIGYVNPTE
 AFIOGYPDOSARRROKOVIEBLADHYLLICGGRGCOOIAPEFAVENIDPVYDASPE
 VTIOAKINDYAVLODADTLDLLELAHIERAICYSALSSDENETYSKATLPKI
 RAIRASSSEAVOKLRAGADEVSPYITGSKRLAALRPQVVSFVGITIGADRSF
 YMEERIGADEDPYIGQTLREAOQLAAGSALITAIRQDRKLIVPMQDTHLLDLSL
 ICLGVEQLRAVNLQCLPLNPAVRALPRNHR"
 10183..11472
 /gene="argd"
 /note="ORF_ID:s1r1022"
 /codon_start=1
 /product="N-acetylornithine aminotransferase"
 /db_xref="PID:g1652236"
 /transl_table=11
 /translation="MTVSPPVSVENAOAFVYNDLSPAAEFKRDEPDVTVMYGRFPI
 AIAKQSTLMDTEKSYLDYVAGIATCTTGHAFALYRAVSDQIKLHYNSLYIP
 EOGELAKNIVESHSCADRVFCNSGAEANEAAIKVRYAHTVLDLEQPVILTAKASF
 HGRITLAIITAGQPYQYQYFDPVLPYGFYVYNDIRSLSENKRYADLEBNSRYAATLE
 PLOGSGVYRPGDLAFKFRVREICDNDILVFEDEVQVGRGKLMGEHGEEDIF
 TSAGLAGVSPIGAMCKKCFDVEPRGNHASFPGNPLAACAGLAVLKTIGDRLLDN
 VOAREGQIRSGIAEIKKQNPILTFEPRWGLINGELISBESSIVSEIVKAMEGGL
 LAPAGPKVLRVYPLVTVTEAETLAQVETLRQATITLV"
 11599..11913
 /note="ORF_ID:s1r1023"
 /codon_start=1
 /product="hypothetical protein"
 /db_xref="PID:g1652237"
 /transl_table=11
 /translation="MKNIMQDRLPDQGFLLLENKNTLERIISGQITPPGEWY
 DQDREPVVLLQGRKAETIYDGAVALISQGDYLLIPAROKHRTVFTTPEPCIMWLALH
 FS"
 11920..12561

CDS
 Note: remainder of annotations omitted.
 Query Match 1.6%; Score 24; DB 20; Length 150894;
 Best Local Similarity 81.6%; Pred. No. 1.38e+00;
 Matches 31; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

DB 4235 aaactctgatatggagccatgcatcatcctcccca 4272
 ||||| ||||| ||||| ||||| ||||| |||||
 QY 1388 AACACTATGATGTGGAGAAAGCACTACATCTCTCCAA 1425

RESULT 15
 LOCUS ECACEA 1344 bp DNA BCT 12-SEP-1993
 DEFINITION E. coli acea gene coding for isocitrate lyase (EC 4.1.3.1).
 ACCESSION X07543
 NID 940884
 KEYWORDS acea gene; isocitrate lyase; isocitritase.
 SOURCE Escherichia coli.
 ORGANISM Escherichia coli.
 Eubacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.

REFERENCE
 AUTHORS Cozzone A.J.
 TITLE Direct Submission
 JOURNAL Submitted (06-MAY-1988) Cozzone A.J., University of Lyon,
 Laboratory of Molecular Biology, 43 Boulevard du onze Novembre,
 69622 Villeurbanne, France

REFERENCE
 AUTHORS Rioult C., Bleicher F., Duclos B., Cortay J.C. and Cozzone A.J.
 TITLE Nucleotide sequence of the acea gene coding for isocitrate lyase in
 Escherichia coli
 JOURNAL Nucleic Acids Res. 16 (12), 5689 (1988)
 MEDLINE 88262573
 FEATURES
 source Location/Qualifiers
 1..1344

THIS PAGE BLANK (USPTO)